

Randomly coloring planar graphs with fewer colors than the maximum degree

Thomas P. Hayes ^{*} Juan C. Vera [†] Eric Vigoda[†]

June 11, 2007

Abstract

We study Markov chains for randomly sampling k -colorings of a graph with maximum degree Δ . Our main result is a polynomial upper bound on the mixing time of the single-site update chain known as the Glauber dynamics for planar graphs when $k = \Omega(\Delta / \log \Delta)$. Our results can be partially extended to the more general case where the maximum eigenvalue of the adjacency matrix of the graph is at most $\Delta^{1-\epsilon}$, for fixed $\epsilon > 0$.

The main challenge when $k \leq \Delta + 1$ is the possibility of “frozen” vertices, that is, vertices for which only one color is possible, conditioned on the colors of its neighbors. Indeed, when $\Delta = O(1)$, even a typical coloring can have a constant fraction of the vertices frozen. Our proofs rely on recent advances in techniques for bounding mixing time using “local uniformity” properties.

^{*}Toyota Technological Institute at Chicago. Email: hayest@cs.uchicago.edu.

[†]College of Computing, Georgia Institute of Technology, Atlanta GA 30332. Email: {jvera,vigoda}@cc.gatech.edu. Research supported in part by NSF grant CCF-0455666

1 Introduction

Markov chains for randomly sampling (and approximately counting) k -colorings of an input graph have been studied intensively in recent years. The colorings problem is appealing as a natural combinatorial problem, as a noteworthy example of a $\#P$ -complete problem, and as a challenging example of the general class of spin systems from statistical physics, which includes problems such as the independent sets (or hard-core model) and Ising model. Improved results for sampling/counting colorings have been in lock-step with advances in the use of coupling techniques. The study of the convergence rate of Markov chains for spin systems has close intuitive (and some formal) connections with macroscopic properties of corresponding statistical physics models.

Considerable attention has been paid to the Glauber dynamics, which is of particular interest for its simplicity and intimate connections to properties of infinite-volume Gibbs distributions (e.g., see [20, 7, 18]). In the Glauber dynamics, at each step, a random vertex is recolored with a color chosen randomly from those colors not appearing in its neighborhood. For a graph with maximum degree Δ , when $k \geq \Delta + 2$ the Glauber dynamics is ergodic with unique stationary distribution uniform over the k -colorings of G . The mixing time of the dynamics is the number of steps, from the worst initial state, to get within variation distance $\leq 1/4$ of the stationary distribution.

A large body of work has studied the following folklore conjecture: For an input graph with maximum degree Δ , the Glauber dynamics has $O(n \log n)$ mixing time whenever $k \geq \Delta + 2$. Such a mixing time is optimal, as shown by Hayes and Sinclair [13], and leads to a fully-polynomial randomized approximation scheme for counting k -colorings for any $k \geq \Delta + 2$. For general graphs, $\Delta + 2$ is a clear lower bound since there exist graphs where the Glauber dynamics is not ergodic below this threshold (and some graphs are not colorable below $\Delta + 1$). We will prove optimal mixing of the Glauber dynamics for $k \ll \Delta$ for a large class of graphs, including all planar graphs.

Martinelli, Sinclair and Weitz [19] proved $O(n \log n)$ mixing time of the Glauber dynamics when $k \geq \Delta + 2$ for the complete $(\Delta - 1)$ -ary tree with arbitrary boundary conditions (that is, a fixed coloring of the leaves). Their result is optimal for worst-case boundary conditions since below $\Delta + 2$ some boundary conditions can “freeze” the entire tree. For graphs with sufficiently large girth $g > 10$ and large maximum degree $\Delta = \Omega(\log n)$, Hayes and Vigoda proved $O(n \log n)$ mixing time when $k \geq (1 + \epsilon)\Delta$, for any $\epsilon > 0$. Their work built upon upon a long series of earlier works (see [9] for a survey), and still seems far from addressing the conjecture without additional girth and degree assumptions. Recently, Hayes [12] presented a relatively simple proof of $O(n \log n)$ mixing time of the Glauber dynamics for any planar graph when $k \geq \Delta + O(\sqrt{\Delta})$.

The $k \geq \Delta + 2$ threshold is a natural threshold from a statistical physics perspective. On the infinite $(\Delta - 1)$ -ary tree, $\Delta + 2$ is the threshold for the persistence of long-range interactions, more precisely, uniqueness/non-uniqueness of infinite-volume Gibbs measures [17, 2]. More precisely, when $k < \Delta + 2$ a fixed coloring of the leaves influences the coloring of the root. In fact, some colorings of the leaves “freeze” the coloring for the remainder of the tree. The existence of frozen colorings on the tree when $k < \Delta + 2$ hints at the major obstacle we need to overcome to prove rapid mixing when $k \ll \Delta$.

In this paper, we get below the $\Delta + 2$ threshold for trees and for all planar graphs. Our

results suggest that for planar graphs the threshold for rapid mixing of the Glauber dynamics is $k = \Theta(\Delta/\log \Delta)$. Note that, even for planar graphs, $\Delta/\log \Delta$ cannot be replaced by a smaller power of Δ , since on any tree of maximum degree Δ , an easy conductance argument shows that the Glauber dynamics has mixing time $\Omega(n \exp(\Delta/k))$, which is superpolynomial in n when $k = o(\Delta/\log n)$. The only previous rapid mixing results for $k < \Delta$ were for 3-colorings of finite subregions of the 2-dimensional integer lattice [11, 10], and random graphs [5].

Our work builds upon the ideas of Hayes [12] to utilize small operator norm ρ . (The operator norm, or “spectral radius,” equals the maximum eigenvalue of any adjacency matrix of the graph.) In addition to the spectral properties, an important component of our work is proving “local uniformity” properties for graphs with small operator norm. For example, showing that for a random coloring, the colors appearing in the neighborhood of a vertex are roughly independent. Such local uniformity properties have been the basis for many previous results for colorings, beginning with [4] (see [9] for a survey). The challenging aspect in our work is that since $k \ll \Delta$, there are nearly frozen colorings, hence even ergodicity is not obvious. For graphs with large maximum degree we prove that the local uniformity properties hold with high probability, building upon [8]. This leads to the following theorem, whose proof uses the coupling with stationarity approach of [14].

Theorem 1.1. *For all $\epsilon > 0$, for all G with operator norm $\leq \Delta^{1-\epsilon}$ and $\Delta = \Omega(\log^{1+\epsilon} n)$, all $k > 4\epsilon^{-1}\Delta/\ln \Delta$, the Glauber dynamics has mixing time $O(n \log n)$.*

Removing the degree restriction presents major obstacles since for a random coloring, a constant fraction of the vertices are frozen. We introduce a new Markov chain which is a more natural chain to both implement and analyze for graphs with operator norm $\leq \Delta^{1-\epsilon}$, $\epsilon > 0$. It is a generalization of the standard dynamics for bipartite graphs in which we alternately recolor all of the vertices in one of the two partitions. We refer to the new chain as the *set dynamics*. We partition the vertices into level sets and then successively recolor the sets.

Consider a partition of the vertices $V = L_0 \cup L_1 \cup \dots \cup L_{m-1}$. The dynamics works in rounds, where in round i we do $|L_j| \log \Delta$ random recolorings (i.e., Glauber updates) of vertices in L_j where $j = i \bmod m$. (If the set L_j is an independent set, then we can instead simply recolor the vertices of L_j once in arbitrary order.) We define the mixing time of the set dynamics as the total number of vertex recolorings until we are within variation distance $\leq 1/4$ of the uniform distribution. Set dynamics can be viewed as a hybrid of Glauber dynamics, which corresponds to the trivial partition, and *systematic scan* dynamics, which correspond to the complete partitions into singletons. Systematic scan is popular in experimental work, but often appears more difficult to analyze than the Glauber dynamics, see, e.g., [3].

We use the set dynamics where the vertices are partitioned into level sets based on their entry in the principal eigenvector. We formally define our partition into level sets in Section 4. We now formally state the main theorem of this paper.

Theorem 1.2. *There exists $C > 0$ such that for every planar graph G , all $k > C\Delta/\log \Delta$, where Δ denotes the maximum degree of G , the following hold:*

- (i) *We can efficiently compute a partition of V into $\ell = O(\log n)$ sets such that the set dynamics mixes within $O(n \log^2 n \log \Delta)$ vertex updates, and*

(ii) *The Glauber dynamics has polynomial mixing time.*

The polynomial mixing time of the Glauber dynamics follows immediately from the above result for the set dynamics. In particular, $O^*(n^4)$ mixing time follows from a straightforward comparison argument as in the proof of Theorem 32 of Dyer et al [3]. The proof of Theorem 1.2 uses the ideas presented in [6] for utilizing local uniformity properties for constant degree graphs.

Although uniqueness of the infinite-volume Gibbs measure may be the key concept for rapid mixing of the Glauber dynamics on general graphs, our results show that, at least in the case of planar graphs, there is a second threshold for rapid mixing. This threshold may correspond to extremality of the free measure (that is, no boundary condition) in the set of infinite-volume Gibbs measures; see [19] and [12] for evidence from the hard-core model on independent sets.

2 Preliminaries

2.1 Basic Notation

We begin by specifying some notation which will be used throughout the paper. Let $G = (V, E)$ be the graph to be colored, and let k denote the number of colors to be used. We say a function $f : V \rightarrow \{1, \dots, k\}$ is a *proper k -coloring* of V if, for every edge $\{u, v\} \in E$, $f(u) \neq f(v)$. Let Ω denote the set of all proper k -colorings of G . For $X \in \Omega$ and $v \in V$, let

$$\mathcal{A}_X(v) = [k] \setminus X(N(v))$$

be the set of available colors for v in X , and let $a_X(v) = |\mathcal{A}_X(v)|$.

2.2 Operator norm and hereditary average degree

Let $\delta = \delta(G)$ denote the maximum, over subgraphs H of G , of the average vertex degree in H . The quantities δ and ρ are closely related as we recall in Section 2, e.g., $\delta \leq \rho$, but they can be significantly different, such as for planar graphs where $\delta < 6$ and $\rho = O(\Delta^{1/2})$.

Theorem 1.1 applies to graphs with $\rho \leq \Delta^{1-\epsilon}$ for any $\epsilon > 0$. Examples of such graphs are the following (e.g., see [1]):

- Planar graphs.
- Graphs embeddable on any fixed surface of finite genus.
- Bipartite graph $G = (V_1 \cup V_2, E)$ has $\rho = \sqrt{\Delta_1 \Delta_2}$ where Δ_i is the maximum degree of vertices in V_i , $i = 1, 2$. Thus bipartite graphs where one side of the bipartition has maximum degree $\Delta^{1-2\epsilon}$ satisfy the assumptions of our theorem.
- Generalizing the previous example, any graph such that the product of degrees of any two adjacent vertices is at most $\Delta^{2-2\epsilon}$.
- Unions of any fixed number of the above, since the operator norm is subadditive.

We will use both δ and ρ in our proofs.

Proposition 2.1.

$$\delta \leq \rho \leq 2\sqrt{\delta\Delta}$$

Proof. Let e_U the characteristic vector of U in V . Then, $|E(U)| = \sum_{u \in U} |N(u) \cap U| = e_U^T A e_U \leq \rho(G) \|e_U\|^2 = \rho(G) |U|$. Thus, average degree in U is $\leq \rho$ which implies $\delta \leq \rho$. The other inequality follows from [12, Theorem 16]. \square

Most of our proofs work with ρ and require that $\rho \leq \Delta^{1-\epsilon}$ for some $\epsilon > 0$. Note, this is equivalent to requiring that $\delta \leq \Delta^{1-\epsilon}$ for some $\epsilon > 0$.

We now point out that, for graphs with small operator norm or, equivalently, small δ , the Glauber dynamics is ergodic with many fewer colors than the maximum degree.

2.3 Upper bounds on diameter of the Glauber dynamics

In general, when $k \leq \Delta + 1$, it is possible that the Glauber dynamics is not connected; for example, when G is the complete graph on $n = \Delta + 1$ vertices. In this case, using $k = \Delta + 1$ colors, every coloring is “frozen,” meaning that the connected components of Ω under the Glauber dynamics are all singletons. However, we restrict our attention to a “nicer” class of graphs, for which we will see that fewer colors are needed.

We now derive some fairly straightforward bounds on the diameter of Ω in terms of the max-min degree over subgraphs of G .

Theorem 2.2. *Suppose every subgraph of G contains at least one vertex of degree $\leq d$. Then for every $k \geq 2(d + 1)$, the Glauber dynamics is ergodic. Indeed, the diameter of Ω with respect to Glauber dynamics is at most $n^2 - n$.*

Proof. Inductively order $V = \{v_1, \dots, v_n\}$ so that, for every i , v_i has at most d neighbors among v_{i+1}, \dots, v_n . This can be done greedily, using the definition of d . Observe that G can now be $(d + 1)$ -colored by simply greedily assigning legal colors to the vertices in the order v_n, v_{n-1}, \dots, v_1 . Fix such a $(d + 1)$ -coloring X . To walk from an arbitrary $Y \in \Omega$ to X , we will proceed in n rounds, as follows.

In round $i \geq 1$, recolor vertices v_i, v_{i-1}, \dots, v_1 in that order. When i has the same parity as n , use only colors from the set $\{1, \dots, d + 1\}$. When i has the opposite parity from n , use colors from the set $\{d + 2, \dots, 2d + 2\}$. Note that there always is such a color available at each step, since whenever recoloring a vertex v_j , at most d colors are forbidden due to neighbors v_i where $i > j$, and no colors from the allowed set of colors (either $\{1, \dots, d + 1\}$ or $\{d + 2, \dots, 2d + 2\}$) are present among the vertices v_i where $i \leq j$, since all were recolored using the other color set in the previous round. In the final round n , choose the colors to agree with X , instead of arbitrarily; since none of these colors are in use at the beginning of round n , there is nothing to prevent this.

Combining this with the triangle inequality shows that the diameter of Ω is at most $2\binom{n}{2} = n^2 - n$. \square

This yields the following two corollaries.

Corollary 2.3. *For planar graphs, when $k \geq 12$, the diameter of Ω is at most $n^2 - n$.*

Proof. We use the fact that a planar graph on n vertices has average degree at most $6(1-2/n)$, which is in turn a consequence of Euler's formula. The result now follows by Theorem 2.2 \square

Corollary 2.4. *For a graph with spectral radius ρ , when $k \geq 2(\rho + 1)$, the diameter of Ω is at most $n^2 - n$.*

Proof. This follows by combining Proposition 2.1 with Theorem 2.2. \square

Next, we show that with a few more colors, the diameter becomes nearly as small as possible.

Theorem 2.5. *For every $k \geq 2\delta + 2$, the diameter of Ω with respect to Glauber dynamics is at most $n \log_{(k-1)/(2\delta)}(n)$.*

Proof. This time we inductively partition V into disjoint subsets S_1, S_2, \dots, S_ℓ , such that for each $v \in S_i$, v has at most $k/2 - 1$ neighbors in $\bigcup_{j \geq i} S_j$. If we do this greedily, we find that for each i ,

$$|S_i| \geq \left(1 - \frac{2d}{k-1}\right) \left|\bigcup_{j \geq i} S_j\right|.$$

Hence $\ell \leq \log_{(k-1)/(2d)}(n)$.

Finally, we now mimic the idea of the proof of Theorem 2.2, except that in round i , we recolor all the vertices of sets S_i, S_{i-1}, \dots, S_1 in that order, using either the colors $1, \dots, \lfloor k/2 \rfloor$ or $\lfloor k/2 \rfloor + 1, \dots, k$, according to parity. Since each round involves at most n recolorings, the diameter bound follows. \square

We remark that, trivially, the diameter of Ω is always at least n , and that $|\Omega| \geq (k-d)^n$, where d is the max-min degree of subgraphs of G . So our diameter bounds are nearly the best possible, which might give some hope that the graph on Ω is a good expander.

2.4 Level Sets

An important component of our work is the partition of G into level sets based on the principal eigenvector of the graph. Let ρ be the operator norm of A , an adjacency matrix of G . Let J denote the $n \times n$ all-ones matrix. Note that the perturbed adjacency matrix $\tilde{A} := A + \frac{\rho}{n}J$ has maximum eigenvalue $\tilde{\rho}$ which satisfies $\rho < \tilde{\rho} \leq 2\rho$.

Let $w \in R_+^n$ be an eigenvector of \tilde{A} such that $\tilde{A}w = \tilde{\rho}w$. Note that this implies that all entries of w are strictly positive. Moreover, $\tilde{\rho}w_v \geq \frac{\rho}{n}\|w\|_1$ for all $v \in V$. In particular, if we let w_{\min} denote the minimum entry of w , then $w_{\min} \geq \frac{1}{2n}\|w\|_1$. Our proofs will analyze a coupling argument using a weighted Hamming distance defined using this eigenvector w . For every $S \subseteq V$ denote $w(S) := \sum_{s \in S} w(s)$. Notice that for every $u \in V$,

$$w(N(u)) = (Aw)(u) \leq (\tilde{A}w)(u) \leq \tilde{\rho}w(u).$$

Our set dynamics for Theorem 1.2 uses level sets defined by w . Let $\epsilon > 0$ be such that $\rho \leq \Delta^{1-\epsilon}$. We define the *level sets*:

$$L_i = \left\{v : \Delta^{i\epsilon/2} \leq \frac{w(v)}{w_{\min}} < \Delta^{(i+1)\epsilon/2}\right\}.$$

Let $L_{<i} = \bigcup_{j < i} L_j$. The level sets are also used for the uniformity results needed in the proof of both theorems.

2.5 Basic Properties of the Level Sets

We first define a bound showing that most neighbors of a vertex lie in lower levels. For $v \in L_i$,

$$w(N(v)) \leq \rho w(v) \leq \Delta^{1-\epsilon} \Delta^{\epsilon(i+1)/2}.$$

Also,

$$w(N(v)) \geq w(N(v) \setminus L_{<i}) \geq |N(v) \setminus L_{<i}| \Delta^{\epsilon i/2}.$$

Hence, for $v \in L_i$,

$$|N(v) \setminus L_{<i}| < \Delta^{1-\epsilon/2}. \quad (1)$$

We now bound the total number of levels. Let m denote the number of levels, and let w_{\max} denote the maximum entry of w . Note,

$$w_{\min} \geq \frac{1}{2n} \|w\|_1 \geq \frac{1}{2n} w_{\max}.$$

Hence,

$$m \leq \frac{2 \ln(2n)}{\epsilon \ln \Delta}. \quad (2)$$

3 Graphs of large degree

In this section, we prove Theorem 1.1 via a coupling argument. Consider two copies of the Glauber dynamics, $(X_t, t \geq 0)$ and $(Y_t, t \geq 0)$. Define the set of disagreements at time t as

$$D_t = \{v \in V : X(v) \neq Y(v)\}.$$

We couple the two processes using Jerrum's coupling [15], i.e., at every step we choose the same vertex in each process and then the choice of colors is maximally coupled. Let $\epsilon > 0$ be such that $\rho \leq \Delta^{1-\epsilon}$ and $k > 4\epsilon^{-1} \Delta / \ln \Delta$. We will prove that if $\Delta = \Omega(\ln^{1+\epsilon} n)$, under Jerrum's coupling we have $\mathbb{E}[w(D_T)] \leq w_{\min}/2$ for $T = O(n \ln n)$, thereby proving Theorem 1.1

For $Y \in \Omega$ and $v \in V$, recall $\mathcal{A}_Y(v) = [k] \setminus Y(N(v))$ and $a_Y(v) = |\mathcal{A}_Y(v)|$. We will use $a_t(v)$ as shorthand for $a_{Y_t}(v)$.

For all $t \geq 0$, given X_t, Y_t , we have

$$\begin{aligned} \mathbb{E}[w(D_{t+1}) | X_t, Y_t] - w(D_t) &= \frac{1}{n} \sum_{v \in V} w(v) \Pr[v \in D_{t+1} \mid v \text{ chosen at time } t] \\ &\quad - \frac{1}{n} \sum_{v \in D_t} w(v) \\ &\leq \frac{1}{n} \sum_{v \in V} w(v) \frac{|N(v) \cap D_t|}{a_t(v)} - \frac{1}{n} w(D_t) \end{aligned} \quad (3)$$

The key to the proof of the theorem will be the following local uniformity property.

Lemma 3.1. *Let $\epsilon > 0$ be given, let G be a graph such that $\rho \leq \Delta^{1-\epsilon}$ and $\Delta = \Omega(\ln^{1+\epsilon} n)$, and let $k > 4\epsilon^{-1}\Delta/\ln \Delta$. Let Y be chosen uniformly from Ω , the set of all proper k -colorings of G . Then,*

$$\Pr [\exists v \in V a_Y(v) < \Delta^{1-\epsilon/2}] \leq n^{-4}. \quad (4)$$

The lemma is related to uniformity properties originally used by Dyer and Frieze [4]. The difficulty in proving the lemma in our setting is that $k \ll \Delta$ and thus we have to consider frozen vertices. Before proving Lemma 3.1, we now use it to complete the proof of Theorem 1.1. The essential point is that colorings for which every vertex has many available colors are universally distance-decreasing, as defined by Hayes and Vigoda in [14]. Since Lemma 3.1 implies that almost all colorings satisfy this property, rapid mixing follows by “coupling with stationarity.” Note that, following Hayes [12], we use a weighted Hamming metric, with weights taken from the principal eigenvector of G .

Assuming Y_0 is chosen uniformly from Ω , Y_t is uniform over Ω . Therefore, conditioning on an event of probability $1 - O(n^{-4})$ we have

$$\begin{aligned} \mathbb{E}[w(D_{t+1})|X_t, Y_t] - w(D_t) &\leq \frac{1}{n\Delta^{1-\epsilon/2}} \sum_{v \in V} w(v)|N(v) \cap D_t| - \frac{1}{n}w(D_t) \\ &= \frac{1}{n\Delta^{1-\epsilon/2}} \sum_{u \in D_t} \sum_{v \in N(u)} w(v) - \frac{1}{n}w(D_t) \\ &\leq \frac{\tilde{\rho}}{n\Delta^{1-\epsilon/2}} \sum_{u \in D_t} w(u) - \frac{1}{n}w(D_t) \\ &\leq -\frac{1}{2n}w(D_t), \end{aligned}$$

since $\tilde{\rho} < 2\rho < \frac{1}{2}\Delta^{1-\epsilon/2}$. Therefore, conditionally

$$\mathbb{E}[w(D_{t+1})] \leq \left(1 - \frac{1}{2n}\right) \mathbb{E}[w(D_t)] \leq e^{-1/2n} \mathbb{E}[w(D_t)].$$

By induction, for $T \geq 2n \ln(2en) \geq 2n \ln(e\|w\|_1/w_{\min})$, we have

$$\mathbb{E}[w(D_T)] \leq e^{-T/2n} \mathbb{E}[w(D_0)] \leq w_{\min}/e,$$

where we have conditioned on an event of probability $1 - O(n^{-3} \ln n)$. Hence, without the conditioning, we have

$$\mathbb{E}[w(D_T)] \leq w_{\min}/e + O(\|w\|_1 n^{-3} \ln n) \leq w_{\min}/2,$$

for sufficiently large n . Markov’s inequality now implies

$$\Pr[w(D_T) \neq 0] < 1/2,$$

which completes the proof of Theorem 1.1. \square

Finally, we prove the uniformity result, Lemma 3.1. In order to deal with the possibility of frozen vertices, we divide the vertices into level sets based on the principal eigenvector.

A simplified example which illustrates the intuition of the proof is the case of the complete $(\Delta - 1)$ -ary tree. To prove the uniformity property we would first consider the leaves which are clearly not frozen. After all of the leaves are recolored, we can consider the parents of the leaves since these vertices are now likely to have some colors available when $k = \Omega(\Delta / \log \Delta)$, and then we continue up the tree by level.

Proof of Lemma 3.1. For $v \in V$ and $Y \in \Omega$, define $\mathcal{G}(Y, v)$ as the event that v has the desired uniformity property under Y , that is,

$$a_Y(v) \geq \frac{1}{2} k e^{-\Delta/k}.$$

Similarly, for $U \subseteq V$, let $\mathcal{G}(Y, U)$ denote the intersection of the events $\mathcal{G}(Y, v)$, for all $v \in U$. We will prove, by induction over levels, that if Y is chosen uniformly in Ω ,

$$\Pr[\neg \mathcal{G}(Y, L_{\leq i})] \leq 2^i |L_{\leq i}| p, \text{ for all } i, \quad (5)$$

where $p = n^{-6}$. It will follow that

$$\Pr[\neg \mathcal{G}(Y, V)] \leq 2^m n p \leq n^2 p \leq n^{-4},$$

where the bound $2^m \leq n$ follows from (2) assuming $\Delta \geq \exp(4/\epsilon)$. The base case $i = 0$ follows vacuously. Now fix $i \geq 0$ and $v \in L_{i+1}$.

All but a few neighbors of v are in previous levels, and all but a few have small co-degree with v . Let S be the set of vertices satisfying both properties. In particular, let

$$S = \{u \in N(v) : u \in L_{\leq i} \text{ and } |N(u) \cap N(v)| \leq \rho \Delta^{\epsilon/2}\}.$$

Let $\bar{S} = N(v) \setminus S$. Notice that, by Proposition 2.1,

$$|\bar{S} \cap L_{\leq i}| \rho \Delta^{\epsilon/2} \leq \sum_{u \in N(v)} |N(u) \cap N(v)| \leq \rho \Delta.$$

Thus,

$$|\bar{S} \cap L_{\leq i}| \leq \Delta^{1-\epsilon/2}.$$

On the other hand, by (1),

$$|\bar{S} \setminus L_{\leq i}| \leq |N(v) \setminus L_{\leq i}| \leq \Delta^{1-\epsilon/2}.$$

Therefore,

$$|\bar{S}| \leq 2\Delta^{1-\epsilon/2}.$$

Thus, all but few of the neighbors of v are in S . We will recolor the vertices in S . Building on the approach used in [8], we will use the small co-degree to show that the colors assigned to S are “fairly independent,” and hence that enough colors remain available for v .

Let $q = |S|$ and write $S = \{s_1, s_2, \dots, s_q\}$. We run the following experiment: Choose $Y \in \Omega$ uniformly at random. Define $Y_0 = Y$ and for each $j = 1, \dots, q$, let $Y_j \in \Omega$ be obtained by recoloring s_j with a color chosen uniformly from $\mathcal{A}_{Y_{j-1}}(s_j)$. We will prove

$$\Pr[\neg \mathcal{G}(Y_q, v) | \mathcal{G}(Y, L_{\leq i})] \leq p. \quad (6)$$

Notice that since $Y_0 = Y$ is uniformly distributed over Ω , so are Y_1, \dots, Y_q . This allows us to deduce

$$\begin{aligned}
\Pr[\neg \mathcal{G}(Y, L_{\leq i+1})] &\leq \Pr[\neg \mathcal{G}(Y, L_{\leq i})] + \Pr[\neg \mathcal{G}(Y, L_{i+1})] \\
&= \Pr[\neg \mathcal{G}(Y, L_{\leq i})] + \Pr[\neg \mathcal{G}(Y_q, L_{i+1})] \\
&\leq 2\Pr[\neg \mathcal{G}(Y, L_{\leq i})] + \sum_{v \in L_{i+1}} \Pr[\neg \mathcal{G}(Y_q, v) | \mathcal{G}(Y, L_{\leq i})] \\
&\leq |L_{\leq i}|2^i p + |L_{i+1}|p \quad \text{by induction and (6)} \\
&\leq |L_{\leq i+1}|2^i p
\end{aligned}$$

To prove (6) we first consider the case in which there are actually no edges between vertices in S . In this case, conditioned on Y , the colors assigned to S under Y_q are fully independent random variables. In the case of the good event $\mathcal{G}(Y, L_{\leq i})$, each color is moreover chosen uniformly from at least $A_{\min} := \frac{1}{2}ke^{-\Delta/k} - \rho\Delta^{\epsilon/2}$ possibilities.

Following Dyer and Frieze [4], this allows us to deduce

$$\begin{aligned}
\mathbb{E}[a_{Y_q}(v) | Y] &\geq \mathbf{1}[\mathcal{G}(Y, L_{\leq i})] \sum_{c \in K} \prod_{j=1}^q \left(1 - \frac{\mathbf{1}[c \in \mathcal{A}_Y(s_j)]}{|\mathcal{A}_Y(s_j)|}\right) \\
&\geq \mathbf{1}[\mathcal{G}(Y, L_{\leq i})] |K| \prod_{c \in K} \prod_{j=1}^q \left(1 - \frac{\mathbf{1}[c \in \mathcal{A}_Y(s_j)]}{|\mathcal{A}_Y(s_j)|}\right)^{1/|K|} \\
&\geq \mathbf{1}[\mathcal{G}(Y, L_{\leq i})] |K| \exp\left(-\frac{1}{|K|} \sum_{c \in K} \sum_{j=1}^q \frac{\mathbf{1}[c \in \mathcal{A}_Y(s_j)]}{|\mathcal{A}_Y(s_j)| - 1}\right) \\
&\geq \mathbf{1}[\mathcal{G}(Y, L_{\leq i})] |K| e^{-qA_{\min}/|K|(A_{\min}-1)} \\
&\geq \mathbf{1}[\mathcal{G}(Y, L_{\leq i})] \frac{9}{10} ke^{-\Delta/k}.
\end{aligned} \tag{7}$$

where $K = [k] \setminus Y(\overline{S})$ denotes the set of colors which could possibly be available to v under Y_q , given Y . Now consider the (Doob) martingale Z_0, \dots, Z_q defined by

$$Z_j = \mathbb{E}[a_{Y_q}(v) | Y, Y_1(s_1), \dots, Y_j(s_j)].$$

Note that $Z_0 = \mathbb{E}[a_{Y_q}(v) | Y]$, while $Z_q = a_{Y_q}(v)$. Because the colors $Y_j(s_j)$ are independent, conditioned on Y , and each step reveals only a single color, it follows that $|Z_j - Z_{j-1}| \leq 1$. Hence the Azuma-Hoeffding inequality yields

$$\begin{aligned}
\Pr\left[a_{Y_q}(v) \leq \frac{8}{10} ke^{-\Delta/k}\right] &\leq \Pr\left[Z_q \leq Z_0 - \frac{1}{10} ke^{-\Delta/k}\right] \\
&\leq \exp\left(-\left(\frac{1}{10} ke^{-\Delta/k}\right)^2 / 2\Delta\right) \\
&\leq p/2
\end{aligned} \tag{8}$$

where in the last step we used the relations

$$ke^{-\Delta/k} \geq k\Delta^{-\epsilon/4} \geq \Delta^{1-\epsilon/3}$$

$$\Delta \geq (\ln n)^{1+\epsilon}$$

and that Δ is sufficiently large as a function of ϵ . This completes the proof of (6) in the case when there are no edges within S .

For the general case, we argue that, assuming $\mathcal{G}(Y, L_{\leq i})$, the edges within S cause a negligible effect on Y_q . To this end, couple the recolorings $Y_0 = Y, Y_1, \dots, Y_q$ on the actual graph with the corresponding recolorings $\tilde{Y}_0 = Y, \tilde{Y}_1, \dots, \tilde{Y}_q$ on the graph with the edges within S deleted. Define the coupling by induction, at each step maximizing the probability that $\tilde{Y}_j(s_j) = Y_j(s_j)$, conditioned on the history.

Now, by the definition of S and because we are assuming the good event $\mathcal{G}(Y, L_{\leq i})$, each update has at most a $\rho\Delta^{\epsilon/2}/A_{\min} \leq \Delta^{-\epsilon/2}$ probability to create a disagreement, regardless of the previous history.

Now by comparison with a sequence of independent coin flips, we see that the probability of having at least $\frac{1}{10}ke^{-\Delta/k}$ disagreements is at most

$$\begin{aligned} \left(\frac{\Delta}{\frac{1}{10}ke^{-\Delta/k}} \right) \Delta^{-\epsilon ke^{-\Delta/k}/2} &\leq \left(\frac{e\Delta}{\frac{1}{10}ke^{-\Delta/k} \Delta^{\epsilon/2}} \right)^{ke^{-\Delta/k}} \\ &\leq \left(\frac{10e}{\Delta^{\epsilon/6}} \right)^{\Delta^{1-\epsilon/3}} \\ &\leq p/2. \end{aligned}$$

This proves (6) in the general case, completing the proof of Lemma 3.1. \square

4 Graphs of low degree

In this section we prove Theorem 1.2.

4.1 Intuition

We first describe the main challenge when Δ is small and try to provide some intuition about how we overcome this obstacle. When the maximum degree is constant, in a random coloring, a constant fraction of the vertices might be frozen. This poses a problem in that the set of disagreeing vertices under our coupling may be highly correlated with the frozen vertices in the two colorings. To see the difficulty, consider the complete $(\Delta - 1)$ -ary tree with a single disagreement at the root v . Suppose all vertices except the leaves have very few available colors (we will later refer to these vertices as nearly frozen). Then in the early stage of the dynamics neighbors of disagreements have few colors available, and thus might have a high probability of becoming a disagreement.

The proof of Lemma 3.1 gives some insight on how to overcome the difficulty of frozen vertices to try to get some sort of independence between the probability that different vertices are frozen. In that proof, a tree-like structure of the graph is exploited recoloring the graph from the “leaves” up. In that way the uniformity property propagates through the tree structure of the graph. By using our set dynamics where the sets correspond to the level sets based on the principal eigenvector we can achieve similar behavior. Once again, vertices will

have the uniformity property with probability roughly $1 - \exp(-\Delta^{1-\epsilon})$, but in this case that means a constant fraction of the vertices will not have the uniformity property. The key is that the graph within a level set is sparse and most neighbors of this set are in earlier sets. Consequently, we will get that vertices within a set are roughly independent of each other, in terms of having the uniformity property.

4.2 Proof Setup

We fix $\epsilon > 0$ where $\rho \leq \Delta^{1-\epsilon}$. We use the level sets as defined in Section 2.4:

$$L_i = \left\{ v : \Delta^{i\epsilon/2} \leq \frac{w(v)}{w_{\min}} < \Delta^{(i+1)\epsilon/2} \right\},$$

where $w_{\min} = \min_{v \in V} w(v)$. Recall, the level set dynamics does $T_i = |L_i| \ln \Delta$ random Glauber steps in L_i .

Consider an arbitrary pair of colorings X_0 and Y_0 , we will analyze Jerrum's coupling for this pair. We separately analyze each round i of the set dynamics. Let X_i and Y_i denote the colorings at the beginning of round i , and let X_i^t and Y_i^t denote the colorings after t steps in level i . Hence, $X_i = X_i^0$ and $X_{i+1} = X_i^T$ where $T = T_i$. When it is clear from the context we may drop the subindex i .

Suppose we have already run the dynamics on levels $0, \dots, i-1$. We will analyze the effect of the updates in level i . Since we have run the dynamics on levels $< i$ we can expect most of the vertices in those sets to have the uniformity properties. Recall, we expect that a vertex has at least $k \exp(-\Delta/k) >> \Delta^{1-\epsilon/4}$ available colors. Hence, let

$$F_i = \{v \in L_i : |[k] \setminus Y_i^0(N(v))| < 2\Delta^{1-\epsilon/4}\}$$

denote the set of *nearly frozen* vertices. Notice that, since every vertex in L_i has at most $\rho\Delta^{\epsilon/2} \leq \Delta^{1-\epsilon/2}$ neighbors in $L_{\geq i}$, we have that for all $t \geq 0$, for $v \in L_i$,

$$v \notin F_i \text{ implies } |\mathcal{A}_t(v)| > \Delta^{1-\epsilon/4}.$$

Therefore, we will consider vertices in $L_i \setminus F_i$ as having the uniformity properties at all times.

We say that \mathcal{H} , a family of graphs, is ϵ -uniform if there is $C > 0$ such that for every graph $G \in \mathcal{H}$ and all $k > C\Delta/\log \Delta$, the following holds. Let Y be a proper k -coloring of G . Then, for all i , after applying the level set dynamics in levels $1, \dots, i-1$ starting with Y , we have that for all $v \in L_i$,

$$\Pr[v \in F_i] \leq e^{-\Delta^{1-\epsilon/3}}.$$

We are going to prove that if family of graphs is ϵ -uniform, then the following generalization of Theorem 1.2(i) holds

Theorem 4.1. *Let $0 < \epsilon < 1$. Let \mathcal{H} be a ϵ -uniform family of graphs. There exists $C > 0$, for all Δ , for every $G \in \mathcal{H}$ with maximum degree Δ and $\rho(G) < \Delta^{1-\epsilon}$, we can efficiently compute a partition of V into $\ell = O(\log n)$ sets such that for all $k > C\Delta/\log \Delta$ the set dynamics mixes within $O(n \log^2 n \log \Delta)$ vertex updates.*

4.3 Analysis: one level of set dynamics

Fix $i > 0$. For $t \geq 0$, let D^t denote the set of disagreements at time t , i.e., $D^t = \{v \in V : X_i^t(v) \neq Y_i^t(v)\}$. Only set L_i is changing and our focus is on disagreements within this set $D^t \cap L_i$.

We will prove the following lemma which bounds the rate of disagreements spreading.

Lemma 4.2. *Let $0 < \epsilon < 1$. Let \mathcal{H} be a ϵ -uniform family of graphs. Then, there exists $C > 0$ such that for all Δ , for all $k > C\Delta/\log \Delta$, and every graph $G \in \mathcal{H}$ with maximum degree Δ and $\rho(G) < \Delta^{1-\epsilon}$, we have:*

$$\mathbb{E} [w(D^T \cap L_i)] \leq (1 + o_\Delta(1)) \Delta^{-\epsilon/4} w(D^0 \cap N(L_i)).$$

Using the lemma it is relatively straightforward to conclude the main theorem. We first present the proof of Lemma 4.2, and then we complete the proof of Theorem 1.2 using Lemma 4.2.

Before proving the lemma, let us first explain where the quantity $\Delta^{-\epsilon/4}$ arises. The lemma is bounding $w(D^T \cap L_i)$, the disagreements at the end of the round within L_i , in terms of the fixed disagreements which influence L_i , namely $D^0 \cap N(L_i)$. In Jerrum's coupling, when $k > \Delta$, a disagreement can spread to a specific neighbor at rate $1/(k - \Delta)$. Thus, disagreements can expand at rate $\Delta/(k - \Delta)$. Given our weighting of the vertices, we expect that the weight of disagreements would spread at rate $\rho/(k - \Delta)$. Finally, assuming vertices had the uniformity properties, we would expect to replace $(k - \Delta)$ by $k \exp(-\Delta/k)$ which is still meaningful for $k < \Delta$. Our lower bound on k will imply $k \exp(-\Delta/k) \gg \Delta^{1-\epsilon/4}$. Then, we will hope to bound the weighted rate of disagreements spreading by $\rho/k \exp(-\Delta/k) < \Delta^{-\epsilon/4}$.

Proof of Lemma 4.2. Let $G \in \mathcal{H}$ be such that $\rho(G) < \Delta(G)^{1-\epsilon}$. We will consider the effect of each $z \in D^0 \cap N(L_i)$ separately. In particular, for each $z \in D^0 \cap N(L_i)$ we are going to bound the expected number of disagreements created in L_i that have z as their *origin* (as in a standard paths of disagreement argument). We use $D^t(z)$ to refer to the set of disagreements at L_i due to z , and thus we are partitioning $D^t \cap L_i = \bigcup_{z \in D^0 \cap N(L_i)} D^t(z)$.

Because of linearity of expectation, to prove the lemma it is enough to prove the following bound. Let $z \in D^0 \cap N(L_i)$. Then,

$$\mathbb{E} [w(D^T(z))] \leq (1 + o(1)) \Delta^{-\epsilon/4} w(z). \quad (9)$$

To prove (9) we consider the ball $B_r(z)$ of radius $r = \Delta^{1-\epsilon/3} \ln^{-2} \Delta$ around z in L_i , namely

$$B_r(z) = \{v \in L_i : \text{dist}(v, z) \leq r\},$$

where $\text{dist}(v, z)$ is the minimum (unweighted) path length from v to z along vertices in L_i . As in [6], we define the following two events:

- $\mathcal{G}_{\text{unif}}$ denotes the event $F \cap B_r(z) = \emptyset$.
- $\mathcal{G}_{\text{ball}}$ denotes the event $D^t(z) \subseteq B_r(z)$ for all $t < T$.

The event $\mathcal{G}_{\text{unif}}$ says that all vertices in the local ball around z have the uniformity property. The second event $\mathcal{G}_{\text{ball}}$ says that the disagreements due to z are contained in this local ball. We will prove (9) by conditioning on the two events as follows:

$$\begin{aligned} \mathbb{E} [w(D^T(z))] &= \mathbb{E} [w(D^T(z)) \mathbf{1}[\overline{\mathcal{G}_{\text{unif}}}]] \\ &\quad + \mathbb{E} [w(D^T(z)) \mathbf{1}[\mathcal{G}_{\text{unif}} \cap \overline{\mathcal{G}_{\text{ball}}}]] \\ &\quad + \mathbb{E} [w(D^T(z)) \mathbf{1}[\mathcal{G}_{\text{unif}} \cap \mathcal{G}_{\text{ball}}]] \end{aligned}$$

We will separately bound the three terms on the right-hand side.

For the last term on the right-hand side we can assume $\mathcal{G}_{\text{unif}}$ and $\mathcal{G}_{\text{ball}}$ hold. Thus, we can assume the uniformity property and we can extend the proof approach of Theorem 1.1 to show the following claim.

Claim 4.3.

$$\mathbb{E} [w(D^T(z)) \mathbf{1}[\mathcal{G}_{\text{unif}} \cap \mathcal{G}_{\text{ball}}]] \leq \Delta^{-\epsilon/4} w(z). \quad (10)$$

Assuming the claim, to complete the proof of (9) we need to bound the probability that $\mathcal{G}_{\text{unif}}$ or $\mathcal{G}_{\text{ball}}$ does not occur, and then use a rough upper bound on the change in weight in these cases. To bound $\Pr [\overline{\mathcal{G}_{\text{unif}}}]$ we use ϵ -uniformity of \mathcal{H} and $|R_z| \leq \Delta^r$. We get, from the union bound,

$$\begin{aligned} \Pr [\overline{\mathcal{G}_{\text{unif}}}] &\leq \exp (r \ln \Delta - \Delta^{1-\epsilon/3}) \\ &\leq \exp (-\Delta^{1-\epsilon/3} (1 - \ln^{-1} \Delta)). \end{aligned} \quad (11)$$

Now to bound $\Pr [\overline{\mathcal{G}_{\text{ball}}} | \mathcal{G}_{\text{unif}}]$ we apply a standard paths of disagreement argument, which gives:

$$\begin{aligned} \Pr [\overline{\mathcal{G}_{\text{ball}}} | \mathcal{G}_{\text{unif}}] &\leq \Delta \Delta^{(1-\epsilon/2)(r-1)} \binom{T}{r} \left(\frac{1}{|L_i| \Delta^{1-\epsilon/4}} \right)^r \\ &\leq \Delta^{\epsilon/2} \left(\frac{eC \Delta^{-\epsilon/4} \ln \Delta}{r} \right)^r \\ &\leq \exp (-\Delta^{1-\epsilon/3} \ln^2 \Delta), \end{aligned} \quad (12)$$

where we have used that $\rho \leq \Delta^{1-\epsilon}$ implies that the maximum degree in the subgraph induced by L_i in G is not bigger than $\Delta^{1-\epsilon/2}$.

Notice that to become a disagreement at time $t+1$, you have to be a neighbor of D^t and therefore

$$\begin{aligned} \mathbb{E} [w(D^{t+1})] &\leq \mathbb{E} [w(D^t) + w(N(D^t))/n_i] \\ &\leq (1 + \rho/n_i) \mathbb{E} [w(D^t)]. \end{aligned}$$

Thus, we can always take the following quantity M as an upper bound on the weight after T steps:

$$w(D^T(z)) \leq w(z) (1 + \rho/|L_i|)^T \leq w(z) \exp (\Delta^{1-\epsilon} \ln \Delta) =: M.$$

Finally, we can complete the proof of (9) using (11), (12), (4.3) and (10) as follows:

$$\begin{aligned}
\mathbb{E} [w(D^T(z))] &= \mathbb{E} [w(D^T(z)) \mathbf{1}[\overline{\mathcal{G}_{\text{unif}}}]] \\
&\quad + \mathbb{E} [w(D^T(z)) \mathbf{1}[\mathcal{G}_{\text{unif}} \cap \overline{\mathcal{G}_{\text{ball}}}]] \\
&\quad + \mathbb{E} [w(D^T(z)) \mathbf{1}[\mathcal{G}_{\text{unif}} \cap \mathcal{G}_{\text{ball}}]] \\
&\leq M\text{Pr} [\overline{\mathcal{G}_{\text{unif}}}] + M\text{Pr} [\overline{\mathcal{G}_{\text{ball}}} | \mathcal{G}_{\text{unif}}] \\
&\quad + \mathbb{E} [w(D^T(z)) \mathbf{1}[\mathcal{G}_{\text{unif}} \cap \mathcal{G}_{\text{ball}}]] \\
&\leq \exp(-\Delta^{1-\epsilon/2})w(z) + (1 + o_\Delta(1))\Delta^{-\epsilon/4}w(z) \\
&\leq (1 + o_\Delta(1))\Delta^{-\epsilon/4}w(z).
\end{aligned}$$

This completes the proof of (9) which also completes the proof of Lemma 4.2. \square

It remains to prove Claim 4.3.

Proof of Claim 4.3. The proof extends the approach for the high degree case in Section 3. The key is that $\mathcal{G}_{\text{ball}}$ and $\mathcal{G}_{\text{unif}}$ together imply that for all of the Glauber updates involving potential disagreements, the updated vertices have at least $\Delta^{1-\epsilon/4}$ available colors. Thus, as in the high degree case we can assume the desired local uniformity property.

Let $n_i = |L_i|$. Fix $t \geq 0$, given X_t, Y_t , we have:

$$\Pr [v \in D^{t+1}(z) : v \text{ chosen at time } t] \leq \frac{|N(v) \cap (D^t(z) \cup \{z\})|}{a_t(v)},$$

and thus, specializing equation (3) to L_i and $D^{t+1}(z)$ we get

$$\begin{aligned}
\mathbb{E} [w(D^{t+1}(z)) | X_t, Y_t] - w(D^t(z)) &\leq \frac{1}{n_i} \sum_{v \in L_i} w(v) \frac{|N(v) \cap (D^t(z) \cup \{z\})|}{a_t(v)} - \frac{w(D^t(z))}{n_i} \\
&\leq \frac{1}{n_i} \sum_{u \in D^t(z) \cup \{z\}} \sum_{v \in N(u) \cap L_i} \frac{w(v)}{a_t(v)} - \frac{w(D^t(z))}{n_i}
\end{aligned}$$

Let $\mathcal{G}(t)$ denote the event $D^\tau(z) \subseteq S$ for all $\tau < t$ and $\mathcal{G}_{\text{unif}}$ and let $\gamma_t = \mathbf{1}[\mathcal{G}(t)]$ be its indicator. The event $\mathcal{G}(t)$ implies that all neighbors of D^t have at least $\Delta^{1-\epsilon/4}$ available colors. Thus,

$$\begin{aligned}
\gamma_t \sum_{v \in L_i} w(v) \frac{|N(v) \cap (D^t(z) \cup \{z\})|}{a_t(v)} &\leq \gamma_t \Delta^{-(1-\epsilon/4)} \sum_{u \in D^t(z) \cup \{z\}} \sum_{v \in N(u) \cap L_i} w(v) \\
&\leq \gamma_t \Delta^{-(1-\epsilon/4)} \rho w(D^t(z) \cup z) \\
&\leq \gamma_t \Delta^{-\epsilon/4} (w(D^t(z)) + w(z) \mathbf{1}[z \notin L_i])
\end{aligned}$$

We then have,

$$\mathbb{E} [w(D^{t+1}(z)) \gamma_{t+1} | X_t, Y_t] \leq \gamma_t \left(w(D^t(z)) \left(1 - \frac{1 - \Delta^{-\epsilon/4}}{n_i} \right) + \frac{\Delta^{-\epsilon/4} w(z) \mathbf{1}[z \notin L_i]}{n_i} \right)$$

Therefore, taking expectations,

$$\mathbb{E} [w(D^{t+1}(z))\gamma_{t+1}] \leq \mathbb{E} [w(D^t(z))\gamma_t] \left(1 - \frac{1 - \Delta^{-\epsilon/4}}{n_i}\right) + \frac{\Delta^{-\epsilon/4}}{n_i} w(z) \mathbf{1}[z \notin L_{=i}]$$

and thus

$$\begin{aligned} \mathbb{E} [w(D^T(z))\gamma_T] &\leq \mathbb{E} [w(D^0(z))\gamma_0] \left(1 - \frac{1 - \Delta^{-\epsilon/4}}{n_i}\right)^T \\ &\quad + \frac{1}{\Delta^{\epsilon/4} - 1} w(z) \mathbf{1}[z \notin L_{\leq i}] \\ &\leq (1 + o_\Delta(1)) \Delta^{-1} w(z) \mathbf{1}[z \in L_{\leq i}] \\ &\quad + (1 + o_\Delta(1)) \Delta^{-\epsilon/4} w(z) \mathbf{1}[z \notin L_{\leq i}] \\ &\leq (1 + o_\Delta(1)) \Delta^{-\epsilon/4} w(z) \end{aligned}$$

This completes the proof of the claim. \square

4.4 Proof of Theorem 4.1

We now complete the proof of Theorem 4.1 using Lemma 4.2.

Proof of Theorem 4.1. Let $G \in \mathcal{H}$ be such that $\rho(G) < \Delta(G)^{1-\epsilon}$. Recall, X_i and Y_i denote the colorings at the beginning of round i , Let D_i denote the disagreements at the beginning of the round, i.e., $D_i = \{v \in V : X_i(v) \neq Y_i(v)\}$. Let m denote the total number of levels.

As observed in Section 2.5, $w(N(v)) \leq \rho w(v)$. Hence, by the definition of the level sets, the neighbors of a vertex are contained within the neighboring $2/\epsilon$ levels. Therefore,

$$N(L_i) \subseteq \bigcup_{-2/\epsilon \leq \ell - i \leq 2/\epsilon} L_\ell.$$

Combining with Lemma 4.2 we have that if $i \equiv j \pmod m$ then the following holds:

$$\begin{aligned} \mathbb{E} [w(D_{j+1} \cap L_i)] &\leq (1 + o_\Delta(1)) \Delta^{-\epsilon/4} \mathbb{E} [w(D_j \cap N(L_i))] \\ &\leq (1 + o_\Delta(1)) \Delta^{-\epsilon/4} \sum_{-2/\epsilon \leq \ell - i \leq 2/\epsilon} \mathbb{E} [w(D_j \cap L_\ell)]. \end{aligned}$$

Note that if $i \not\equiv j \pmod m$, then

$$\mathbb{E} [w(D_{j+1} \cap L_i)] = \mathbb{E} [w(D_j \cap L_i)].$$

For any $j \geq 0$, let $W_j^{\max} = \max\{\mathbb{E} [w(D_j \cap L_i)] : i = 1, \dots, m\}$. For every $s \geq 0$, it follows by induction on $j \leq m$ that if $i < j$, then

$$\mathbb{E} [w(D_{ms+j} \cap L_i)] \leq (1 + o_\Delta(1)) \frac{4}{\epsilon} \Delta^{-\epsilon/4} W_{ms}^{\max} \leq \frac{1}{2} W_{ms}^{\max},$$

and, if $i \geq j$ we have

$$\mathbb{E} [w(D_{ms+j} \cap L_i)] = \mathbb{E} [w(D_{ms} \cap L_i)].$$

Therefore, for all $s \geq 0$, $W_{m(s+1)}^{\max} \leq \frac{1}{2}W_{ms}^{\max}$. Thus, if

$$s = 1 + \log(m\|w\|_1/w_{\min}) = O(\ln(mn)),$$

then, since $W_0^{\max} \leq \|w\|_1$, we have

$$\mathbb{E}[w(D_{ms})] \leq mW_{ms}^{\max} \leq m2^{-s}W_0^{\max} \leq w_{\min}/2.$$

Since $m = O(\ln n)$, the total number of Glauber moves is $O(n \ln^2 n \ln \Delta)$. \square

5 1/3-uniformity of Planar Graphs

We are going to prove that the class of planar graphs is 1/3-uniform. Notice that this is enough to prove Theorem 1.2 from Theorem 4.1.

The difficulty in proving the uniformity properties for planar graphs when Δ is constant, is that the tail probability is constant and thus a constant fraction of the graph is nearly-frozen in expectation. To prove the desired uniformity property, we need to consider a stronger statement on sets of vertices being nearly-frozen. Roughly speaking, we show that for sets of vertices we can find a large subset such that the probability the vertices are nearly-frozen is roughly independent over the vertices in the subset.

The proof of the uniformity lemma uses the following structural lemma.

Lemma 5.1. *Let G be a graph of maximum degree Δ , let $\delta = \delta(G)$, and let $U \subset V$. Then there exists $U' \subseteq U$ such that*

$$|U'| \geq |U|3^{-\delta}.$$

and for all $u \in U'$,

$$|N(u) \cap N(U' \setminus \{u\})| \leq \delta.$$

Proof. Notice that by the definition of δ we can cover G by the union of δ forests. We can color each forest with 3 colors, say c_1, c_2 and c_3 , such that any vertex has at most one neighbor in common with vertices of the same color. Now partition U according to the colors received by the vertices in the δ colorations and let U' the largest of the parts. \square

Now we introduce some notation. Given v we denote by $\ell(v)$ the level of v , i.e., the ℓ such that $v \in L_\ell$. We define $M = \max_{\{u,v\} \in E(G)} |\ell(v) - \ell(u)|$. Thus for a planar graph $M \leq 4$.

We denote the down-neighbors of v as

$$N^-(v) = \{u \in N(v) : \ell(u) < \ell(v)\},$$

and let $N^+(v) = N(v) \setminus N^-(v)$. We say that u is a descendant of v if there is a path $v = v_1, v_2, \dots, v_m = u$, such that $v_{i+1} \in N^-(v_i)$ for all $i < m$. Finally, we define $D(v)$ the set of descendants of v that are at most M levels down from v .

We will use the following corollary of Lemma 5.1 for planar graphs.

Corollary 5.2. *For any planar graph $G = (V, E)$, any $S \subset V$, there exists $S_1 \subseteq S$ such that $|S_1| \geq |S|3^{-2000}$ and for all $u \in S_1$ we have $|D(u) \cap D(S_1 \setminus \{u\})| < 2000$.*

Proof. Consider the G^M , the graph constructed by joining any two vertices whose distance in G is at most M . Notice that $\delta(G^M) \leq \delta(G)^M \leq 6^4 < 2000$. From Lemma 5.1 we have that there is $S_1 \subseteq S$ such that $|S_1| \geq |S|3^{-2000}$. Moreover if $u \in S_1$ we have $|D(u) \cap D(S_1 \setminus \{u\})| |N_{G^M}(u) \cap \mathcal{N}_{G^M}(S_1 \setminus \{u\})| \leq \delta(G^M)$. \square

We now return to the main proof in this section.

In the following $\eta = \eta(C, \epsilon)$ is a sufficiently small constant. We define a notion of it i -thawed inductively where 0-thawed means the vertex has many available colors. All vertices are $\Delta^{1/4}$ -thawed. We then say $v \in V$ is i -thawed if, for $m = \ell(v)$ we have $|\mathcal{A}_{Y_m}(v)| > 3\Delta^{1-\eta} - i\Delta^{9/10}$ and at most $\Delta^{8/9}$ of its neighbors are not $(i+1)$ -thawed. We denote this set of i -thawed vertices as \mathcal{G}_i .

We now prove the following lemma which is our main uniformity result.

Lemma 5.3. *Let G be a planar graph. Let $S \subseteq L_m$, then*

$$\Pr[S \cap \mathcal{G}_1 = \emptyset] \leq p^{|S|},$$

where $p = e^{-\Delta^{8/9}}$

Note, Lemma 5.3 implies that planar graphs are $1/3$ -uniform.

Proof. We start by assuming $|N^-(v)| \geq \Delta^{1-\eta}$ for all $v \in S$, because for any coloring Y , $|\mathcal{A}_Y(v)| \geq k - |N(v)| \geq K - \delta - |N^-(v)|$.

Let $C = 3^{2000}$ and let $S_1 \subseteq S$ as in Corollary 5.2. We are going to show that the probability of all the elements of S_1 being bad is smaller than $p^{C|S_1|}$.

Let G^* be the graph obtained by deleting for any $u \in D(S_1)$ all the edges between u and $N^+(u)$. Note those vertices in $N(S_1)$ are a subset of $D(S_1)$ and in G^* there are no edges within $N(S_1)$, and from $N(S_1)$ to other vertices in the same or higher level.

The idea of the proof is as in the proof of Lemma 3.1 to run the level set dynamics on G^* where we get independence between vertices in $N(S_1)$, and then to couple the dynamics on G^* with the dynamics on G . Let \mathcal{G}_i^* the set of vertices which are i -thawed when we run the dynamics on G^* . We will prove by induction that most of the neighbors of S_1 are in \mathcal{G}_1^* . We are also going to show that most of the vertices in S_1 have at least $3\Delta^{1-\eta}$ available colors. Therefore, this will show that most of the vertices in S_1 are in \mathcal{G}_0^* . This gives us some slack to show that most of the vertices in S_1 are in \mathcal{G}_1 . To do this, we couple the level set dynamics on G with that on G^* in such a way that for most vertices in $D(S_1)$ the number of disagreements in their neighborhood is bounded by $\Delta^{9/10}$, which suffices to show that most of the vertices in S_1 are 1-thawed in G .

We first use induction to show that for most vertices v in S_1 , most of the neighbors of v are 1-thawed. Let

$$S_2 = \{v \in S_1 : |N^-(v) \setminus \mathcal{G}_1^*| \leq \Delta^{8/9}\}.$$

Notice that

$$|S_1 \setminus S_2|(\Delta^{8/9} - 2000) \leq |N(S_1) \setminus \mathcal{G}_1^*|.$$

Therefore, using the induction hypothesis,

$$\begin{aligned}
\Pr[|S_2| \leq 99|S_1|/100] &\leq \Pr[|N(S_1) \setminus \mathcal{G}_1^*| \leq |S_1|\Delta^{8/9}/101] \\
&\leq \sum_{i=0}^{M-1} \Pr[|N(S_1) \cap L_{\ell(v)-M+i} \setminus \mathcal{G}_1^*| \leq |S_1|\Delta^{8/9}/101M] \\
&\leq \sum_{i=0}^{M-1} \binom{|N(S_1) \cap L_{\ell(v)-M+i} \setminus \mathcal{G}_1^*|}{|S_1|\Delta^{8/9}/101M} p^{|S_1|\Delta^{8/9}/101M} \\
&\leq M \left(\frac{101Me|N(S_1)|}{|S_1|\Delta^{8/9}} p \right)^{|S_1|\Delta^{8/9}/101M} \\
&\leq M (101Me\Delta^{1/4}p)^{|S_1|\Delta^{8/9}/101M} \\
&\leq p^{10C|S_1|}.
\end{aligned} \tag{13}$$

Since most of the neighbors of vertices in S_2 have many available colors, we expect the vertices in S_2 to also have many available colors. We now prove this for the graph G^* . Let

$$S_3 = \{v \in S_2 : |\mathcal{A}_m^*(v)| \geq 3\Delta^{1-\eta}\}.$$

Notice that the elements of S_3 are 0-thawed in G^* . When we run the process on G^* the color choices on $N(S_1)$ are mutually independent and using this we will show that most vertices in S_2 are also in S_3 . Most down-neighbors of S_2 are 1-thawed by definition. Therefore, for $v \in S_2$, at most ρ of its neighbors are not down-neighbors, at most $\Delta^{8/9}$ of its down-neighbors have few (or even no) available colors, and the remaining down-neighbors have at least $\Delta^{1-\eta}$ available colors. By the construction of G^* the color choices on $N(v)$ are mutually independent. Also the probability of not hitting v when running the dynamics on $L_{\ell(v)}$ is smaller than $(1 - 1/n_{\ell(v)})^{n_{\ell(v)} \log \Delta} \leq 1/\Delta$. As in the proof of (7), we have that

$$\mathbb{E}[a_{Y_m^*}(v)] \geq \frac{9}{10}k \exp(-\Delta/k),$$

where Y^* is the chain run on the graph G^* .

Moreover, as in (8), we have that

$$\Pr\left[a_{Y_m^*}(v) \leq \frac{8}{10}ke^{-\Delta/k}\right] \leq p^{20C}. \tag{14}$$

Since, by the construction of G^* , the recolorings of any $w \in N^-(S_2)$ are mutually independent, we have that (14) is independent for $v \in S_2$ and therefore:

$$\Pr[|S_2 \setminus S_3| \geq |S_2|/100] \leq p^{10C|S_2|}. \tag{15}$$

Now we couple the process Y on G with the process Y^* on G^* . For $v \in S_1$, most $w \in N^-(N^-(v))$ are 2-thawed by definition, and continuing in this manner, most $u \in D(S_2)$ are M -thawed. Therefore, for most $u \in D(S_2)$ we have $|\mathcal{A}_{\ell(u)}(u)| \geq 3\Delta^{1-\eta} - M\Delta^{9/10}$ and thus every time u is recolored it has at least $\Delta^{1-\eta}$ colors available. For any $u \in S_2 \cup D(S_2)$,

at most $\Delta^{8/9}$ of its neighbors do not have $\Delta^{1-\eta}$ colors available, and we consider these the bad neighbors of u . Disagreements between G and G^* originate at and propagate through vertices in $D(S_2)$. The lower bound on the available colors for most $u \in D(S_2)$, and the upper bound on the number of bad neighbors, will be used to upper bound the probability of a disagreement propagating.

To couple the processes on G and G^* , we pick the same vertex z in both processes at every step and couple the new color choice for z to maximize the probability that the same color is chosen. Disagreements originate along edges deleted from G , thus they are from some $u \in D(S_2)$ to an up-neighbor. So in some sense disagreements start at vertices $u \in D(S_2)$, they can then spread within the level and up, reaching vertices in $N(S_2)$. For those $u \in D(S_2)$ with many available colors we can upper bound the probability a disagreement reaches u . We say that a disagreement is born at u when it is caused by a disagreement propagating from a bad neighbor or from a neighbor within the level of u . Otherwise, we say the disagreement is propagating (through u). Since u is M -thawed, it has at most $\Delta^{8/9}$ bad neighbors and at most δ other neighbors within its level which can cause a disagreement to be born at u . Thus, the probability of a disagreement being born at u is

$$\leq (\Delta^{8/9} + \delta)/\Delta^{1-\eta} \leq 2\Delta^{-1/9+\eta}.$$

The probability of a disagreement being propagated to u is

$$\leq \Delta^{-1+\eta}.$$

We now want to upper bound the number of disagreements in $N^-(v)$ for each $v \in S_3$, and we will combine this upper bound with (14) to get an upper bound on the number of available colors for v in Y . Let $v \in S_3$ and let H_v be the subgraph of G^* induced by all the good vertices in $D(v)$. In H_v the disagreement between G^* and G is bounded by the process solely on H_v where disagreements are born with probability $2\Delta^{-1/9+\eta}$ and propagated upward through an edge with probability $\Delta^{-1+\eta}$. Thus, at time t if $w \in H_v$, a disagreement is created at w with probability

$$\leq 2\Delta^{-1/9+\eta} + |N^-(w) \cap \mathcal{D}_t| \Delta^{-1+\eta} \quad (16)$$

where \mathcal{D}_t is the set of disagreements at time t . Consider $t = 0$ as the start of recoloring level $\ell(v) - M$, and for $i = 0, \dots, M$, let T_i denote the time we start recoloring level $\ell - M + i$.

We will upper bound the number of disagreements in H_v by induction, starting M levels from v . Let \mathcal{E}_i be the event

$$\forall u \in H_v \cap L_{\ell(v)-M+i}, |N^-(u) \cap \mathcal{D}_{t_u}| \leq \Delta^{8/9+2i\eta}.$$

Thus we allow the number of disagreements to slowly increase as we move closer to v , but it is always bounded by a relatively small quantity. Let

$$\mathcal{E}_{<i} = \cup_{j<i} \mathcal{E}_j.$$

Notice that $\Pr[\mathcal{E}_0] = 0$. We are going to prove for all $i = 1, \dots, M$

$$\Pr[\neg \mathcal{E}_i \mid \mathcal{E}_{<i}] \leq e^{-\Delta^{8/9+\eta/4}} \quad (17)$$

Recall $v \in S_3$ implies v is 0-thawed in G^* . So, if the number of disagreements in $N^-(u)$ is smaller than $\Delta^{9/10}$ for all $u \in H_v$ then $v \in \mathcal{G}_1$. And, assuming (17), observe

$$\begin{aligned} \Pr [\exists u \in H_v, |N^-(u) \cap \mathcal{D}_{t_u}| \geq \Delta^{9/10}] &\leq \Pr [\mathcal{E}_0] + \sum_{i=1}^M \Pr [\mathcal{E}_i \mid \neg \mathcal{E}_{i-1}] \\ &\leq M e^{-\Delta^{8/9+\eta/4}} \\ &\leq e^{-\Delta^{8/9+\eta/8}} \end{aligned}$$

We defer the proof of (17) momentarily, and complete the remaining part of the proof of the Lemma.

Recall Corollary 5.2 If $v, v' \in S_3$ then H_v and $H_{v'}$ are disjoint and therefore the disagreement percolation process in H_v is independent of the disagreement percolation process in $H_{v'}$. So $|S_3 \setminus \mathcal{G}|$ is stochastically bounded by $B \sim \text{Binomial}(|S_3|, e^{-\Delta^{8/9+\eta/8}})$ and therefore

$$\Pr [S_3 \cap \mathcal{G} = \emptyset \mid |S_3|] \leq e^{-|S_3| \Delta^{8/9+\eta/8}}. \quad (18)$$

Thus, putting together (13), (15) and (18) we get

$$\begin{aligned} \Pr [S \cap \mathcal{G} = \emptyset] &\leq \Pr [S_1 \cap \mathcal{G} = \emptyset] \\ &\leq \Pr [S_3 \cap \mathcal{G} = \emptyset \mid |S_3| > 9|S_1|/10] + \Pr [|S_3| < 9|S_1|/10 \mid |S_2| > 99|S_1|/100] \\ &\quad + \Pr [|S_2| < 99|S_1|/100] \\ &\leq e^{-9|S_1| \Delta^{8/9+\eta/8}/10} + p^{10C|S_1|} + p^{10C|S_1|} \\ &\leq p^{C|S_1|} \end{aligned}$$

To finish the proof we only need to prove (17). Note, $H_v \cap L_{\ell(v)-M+i}$ is an independent set in H_v , for every $i = 0, \dots, M$. For $u \in H_v \cap L_{\ell(v)-M+i}$, \mathcal{E}_{i-1} implies $|N^-(w) \cap \mathcal{D}_{t_w}| < \Delta^{8/9+2(i-1)\eta}$ for every $w \in N^-(u)$. The fact that $w \in \mathcal{D}_{t_u}$ depends on whether w is chosen by the dynamics, and if it becomes a disagreement the last time it is chosen. Using (16) we get

$$\Pr [w \in \mathcal{D}_{t_u} \mid \mathcal{E}_{i-1}] \leq 2\Delta^{-1/9+\eta} + \Delta^{-1/9+(2i-1)\eta} \leq \Delta^{-1/9+(2i-1/2)\eta},$$

and thus, conditioning on \mathcal{E}_{i-1} $|N^-(u) \cap \mathcal{D}_{t_u}|$ is dominated by $\text{Binomial}(\Delta, \Delta^{-1/4+(2i-1/2)\eta})$. Using Chernoff bounds we get

$$\Pr [|N^-(u) \cap \mathcal{D}_{t_u}| > \Delta^{8/9+2i\eta} \mid \mathcal{E}_{i-1}] \leq e^{-\Delta^{8/9+(2i+1/2)\eta}} \leq e^{-\Delta^{8/9+\eta/2}}.$$

From the union bound, we have

$$\Pr [\mathcal{E}_i \mid \mathcal{E}_{i-1}] \leq \Delta^M e^{-\Delta^{8/9+\eta/2}} \leq e^{-\Delta^{8/9+\eta/4}}.$$

This completes the proof of (17), and completes the proof of the lemma. \square

6 Concluding Remarks

There are several avenues for further research highlighted by our work. The most immediate direction are better lower-bound examples. For instance, is the mixing time super-polynomial for the Glauber dynamics for the complete $(\Delta - 1)$ -ary tree with $k = 3$, when $\Delta = O(1)$?

Another intriguing direction is proving rapid mixing of the Glauber dynamics for $k < \Delta$ for general bipartite graphs. It is even possible that there are efficient sampling algorithms for triangle-free graphs when $k < \Delta$ since Johansson [16] has shown that the chromatic number of such graphs is $O(\Delta/\log \Delta)$.

References

- [1] N. Biggs. *Algebraic Graph Theory*. 2nd Ed., Cambridge Mathematical Library, Cambridge University Press, 1993.
- [2] G. Brightwell and P. Winkler. Random Colorings of a Cayley Tree. In *Contemporary Combinatorics*. Bolyai Society Mathematical Studies, vol. 10. Bollobás, B. (Ed.), 247-276, 2002.
- [3] M. Dyer, L. A. Goldberg and M. Jerrum. Systematic Scan for Sampling Colorings. *Ann. Applied Prob.*, 18(1):185-230, 2006.
- [4] M. Dyer and A. Frieze. Randomly colouring graphs with lower bounds on girth and maximum degree. *Random Struct. Algorithms*, 23(2):167-179, 2003.
- [5] M. Dyer, A. Flaxman, A. Frieze, and E. Vigoda. Randomly coloring sparse random graphs with fewer colors than the maximum degree. *Random Struct. Algorithms*, 29(4):450-465, 2006.
- [6] M. Dyer, A. Frieze, T. Hayes and E. Vigoda. Randomly coloring constant degree graphs, in *Proceedings of the 45th Annual Symposium on Foundations of Computer Science (FOCS)*, 582-589, 2004.
- [7] M. Dyer, A. Sinclair, E. Vigoda and D. Weitz. Mixing in time and space for lattice spin systems: A combinatorial view. *Random Struct. Algorithms*, 24(4):461-479, 2004.
- [8] A. Frieze and J. Vera. On randomly colouring locally sparse graphs. *Discrete Mathematics and Theoretical Computer Science*, 8(1):121-128, 2006.
- [9] A. Frieze and E. Vigoda. Survey of Markov Chains for Randomly Sampling Colorings. *Combinatorics, Complexity and Chance*, Oxford University Press, 2007.
- [10] L. A. Goldberg, R. Martin and M. Paterson. Random sampling of 3-colorings in Z^2 . *Random Struct. Algorithms*, 24(3):279-302, 2004.
- [11] M. Luby, D. Randall and A. Sinclair. Markov chain algorithms for planar lattice structures. *SIAM J. Comput.*, 31(1):167-192, 2001.

- [12] T. P. Hayes. A simple condition implying rapid mixing of single-site dynamics on spin systems. In *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2006.
- [13] T. P. Hayes and A. Sinclair. A general lower bound for mixing of single site dynamics on graphs. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 511-520, 2005.
- [14] T.P. Hayes and E. Vigoda, Coupling with the Stationary Distribution and Improved Sampling for Colorings and Independent Sets. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005, 971-979.
- [15] M.R. Jerrum, A very simple algorithm for estimating the number of k -colourings of a low-degree graph, *Random Struct. Algorithms*, 7(2):157–165, 1995.
- [16] A. Johansson, Asymptotic choice number for triangle free graphs, *DIMACS Technical Report*, 91-95, 1996.
- [17] J. Jonasson. Uniqueness of uniform random colorings of regular trees. *Statistics & Probability Letters*, 57:243–248, 2002.
- [18] F. Martinelli. *Lectures on Glauber dynamics for Discrete Spin Models*. Lecture Notes in Mathematics, vol. 1717, 2000.
- [19] F. Martinelli, A. Sinclair and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. To appear in *Random Struct. Algorithms*.
- [20] D. Weitz. PhD Thesis. *Mixing in Time and Space for Discrete Spin Systems*. Ph.D. thesis, UC Berkeley, 2004. Available from <http://dimacs.rutgers.edu/~dror/>